

### **REMARKS**

The Examiner rejected claims 1-5, 7-12, 29, 31-33, 42, 43, and 65 as being anticipated by Kawashima et al., Indexing of Baseball Telecast For Content-based Video Retrieval.

Kawashima et al. disclose a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima et al. disclose a first stage that detects domain specific scene in baseball video based-on image similarity, referred to as a basic scene. Each of the basic scene include a pitching and batting action including a significant number of frames on both ends of the pitching and batting action. The exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. Accordingly, Kawashima et al. disclose using a model sequence of frames to determine the approximate frame that the ball is put into play.

Claim 1 patentably distinguishes over Kawashima et al. by claiming an event characterized by a start time based upon when the ball is put into play in a manner free from using a model sequence of frames to determine the approximate frame that the ball is put into play. Using a model sequence of frames, as disclosed by Kawashima et al., to determine the approximate frame that the ball is put into play is generally unreliable and requires significant computational resources.

Claims 2-4 depend from claim 1 and are patentable for the same reasons asserted for claim 1.

Kawashima et al. disclose a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima et al. disclose the identification of basic scenes that include a pitching and batting action including a significant number of frames on both ends of the pitching

and batting action. The exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. Accordingly, Kawashima et al. disclose using basic scenes and dynamic programming to only determine batting sequences.

Claim 5 patentably distinguishes over Kawashima et al. by claiming identifying a plurality of segments of the video based upon a series of activities defined by the rules of baseball that could potentially result in at least one of, (1) a score, (2) preventing a score, (3) advancing a team toward a score, and (4) preventing advancing a team toward a score, where the segments are free from including a batter swinging a bat. For example, these claimed aspects of baseball may include stolen bases or throwing out a runner.

Claims 6-8 depend from claim 5 and are patentable for the same reasons asserted for claim 5.

Kawashima et al. disclose a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima et al. disclose a first stage that detects domain specific scene in baseball video based-on image similarity, referred to as a basic scene. Each of the basic scene include a pitching and batting action including a significant number of frames on both ends of the pitching and batting action. The exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. Accordingly, Kawashima et al. disclose using a model sequence of frames to determine the approximate frame that the ball is put into play.

Claim 9 patentably distinguishes over Kawashima et al. by identifying a plurality of segments of the video based upon detecting a play of the baseball game, wherein the identifying

includes detecting the start of the play in a manner free from using a model sequence of frames to determine the approximate frame that a ball is put into play.

Claims 10-12 depend from claim 9 and are patentable for the same reasons asserted for claim 9.

Claim 29 patentably distinguishes over Kawashima et al. by identifying for the start of at least one of the segments is free from using a model sequence of frames to determine the approximate frame that the ball is put into play.

Claims 30-32 depend from claim 29 and are patentable for the same reasons asserted for claim 29.

Kawashima et al. disclose a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima et al. disclose a first stage that detects a basic scene. The exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image. Thereafter, the end of the segments is identified from a rough estimate of the viewing angle of the next cut just after the pitching. See, Kawashima et al., 2.1.4. Accordingly, after identifying the start of the segment using the model sequence of frames the next scene cut is used to identify the end.

Claim 33 patentably distinguishes over Kawashima et al. by claiming identifying a plurality of segments of the baseball video, wherein the start of the segment is first identified for at least one of the segments then wherein the identifying for the end of at least one of the segments is based upon detecting a plurality of scene changes, where the first scene change detected after the start of the at least one of the segments is free from being selected as the end of the at least one of the segments.

Claims 34-35 depend from claim 33 and are patentable for the same reasons asserted for claim 33.

Kawashima et al. disclose that the model sequence of frames use continuous dynamic programming matching for fixed areas of the image. In particular, Kawashima et al. disclose using areas that are fixed around the bat and around the arms of the pitcher. See 2.1.3, Kawashima et al. last paragraph.

Claim 42 patentably distinguishes over Kawashima et al. by claiming where the processing the processing characterizes whether a batter is sufficiently close to at least one of a catcher and an umpire proximate home base. Kawashima et al. fails to include such a characterization.

Claim 43 depends from claim 42 and is patentable for the same reasons asserted for claim 42.

Kawashima et al. at most describes the storing of clips in a database that are retrievable from a browsing system by specifying the type of clip {hit, homer, struck out}, innings, score, and count. Accordingly, the clips are organized in some manner without regard to operation with other data systems and sources.

Claim 65 patentably distinguishes over Kawashima et al. providing a TV-Anytime compliant description scheme of said video. There is no suggestion in Kawashima et al. that there would be any benefit from incorporating a TV-Anytime compliant description scheme.

The Examiner rejected claims 47, 48, 54, 59, and 60-64 as being anticipated by Rui et al., Automatically Extracting Highlights for TV Baseball Programs.

Rui et al. disclose a system for extracting highlights for TV baseball programs. The focus of Rui et al. is on audio-only highlight extraction. See, Rui et al., page 105.

Claim 47 patentably distinguishes over Rui et al. by claiming creating a summarization of the video based upon analysis of the video by including said plurality of segments and removing at least one of the segments from the summary based, at least in part, upon audio information related to the at least one of the segments.

Claim 48 depends from claim 47 and is patentable for the same reasons asserted for claim 47.

Rui et al. describes the use in indices and descriptions. Accordingly, there is no regard to operation with other data systems and sources.

Claim 54 patentably distinguishes over Rui et al. by claiming providing a TV-Anytime compliant description scheme of said video and identifying a plurality of segments of said video based upon said TV-Anytime compliant description scheme.

Claims 55-56 depend from claim 54 and are patentable for the same reasons asserted for claim 54.

Rui et al. disclose a system for extracting highlights. The applicant respectfully suggests that there is no service provider, description, and remote user. The Examiner points to "a video summary is transmitted in response to user's request via indices/descriptions" on page 105. The applicant cannot locate such a comment in Rui et al. In any event, Rui et al. fails to suggest the use of summaries where the user is remote from a service provider.

Claim 59 patentably distinguishes over Rui et al. by claiming a service provider identifying a summary of a video, a service provider identifying a description of said summary,

and said service provider providing said summary and said description to a remote user in response to receiving a request from said remote user.

Claim 60 depends from claim 59 and is patentable for the same reason asserted for claim 59.

Claim 61 patentably distinguishes over Rui et al. by claiming a service provider storing a summary of a video that includes baseball and the service provider providing the summary in the form of streaming video to a user that is interconnected to the service provider by a network, in response to receiving a request from the user.

Claim 62 depends from claim 61 and is patentable for the same reason asserted for claim 61.

Claim 63 patentably distinguishes over Rui et al. by claiming a service provider storing a description of a summary of a video that includes baseball and the service provider providing the description to a user that is interconnected to the service provider by a network in response to receiving a request from the user

Claim 64 depends from claim 63 and is patentable for the same reason asserted for claim 63.

The Examiner rejected claims 13-28, 30, 35, 39, 44, 45, 51-53 as being unpatentable over Kawashima et al. in view of Wang et al., Multimedia Content Analysis.

Kawashima et al. disclose a system for indexing baseball telecast for content-based video retrieval. In particular, Kawashima et al. disclose a first stage that detects domain specific scene in baseball video based-on image similarity, referred to as a basic scene. The first stage is determined based upon using a similarity measure between an input image and a model image,

together with a threshold. In a second stage, the exact location of pitching and batting actions is determined using continuous dynamic programming matching for fixed areas of the image.

Accordingly, Kawashima et al. disclose using a similarity measure between a pair of frames and a model sequence of frames to determine the approximate frame that the ball is put into play, again based upon a set of model sequence. In essence, Kawashima et al. teach a technique based upon interpreting the video containing baseball in order to identify portions of interest.

It is possible to develop highly sophisticated models of a typical baseball video to identify potentially relevant portions of the video, as taught by Kawashima et al. However, such highly sophisticated models are difficult to create and are not normally robust. Further, the likelihood that a majority of the highly relevant portions of the baseball video will be included in such a video summarization is low because of the selectivity of the model. Thus the resulting video summarization of the baseball game may simply be unsatisfactory to the average viewer.

After consideration of the difficulty of developing highly sophisticated models of a baseball video to analyze the content of the baseball video, as the sole basis upon which to create a baseball summarization, the present inventors determined that this technique is ultimately flawed as the models will likely never be sufficiently robust to detect all the desirable content. Moreover, the number of different types of model sequences of potentially desirable content is difficult to quantify. In contrast to attempting to detect particular model sequences, the present inventors determined that the desirable segments of the baseball game are preferably selected based upon a "play." A "play" may be defined as an sequence of events defined by the rules of baseball. In particular, the sequence of events of a "play" may be defined as the time generally at which the ball is put into play (e.g., a time based upon when the ball is put into play) and the

time generally at which when the ball is considered out of play (e.g., a time based upon when the ball is considered out of play). Normally the “play” would include a related series of activities that could potentially result in a score (or a related series of activities that could prevent a score) and/or otherwise advancing the team toward scoring (or prevent advancing the team toward scoring).

The present inventors then considered how to detect a “play” from a baseball video in a robust, efficient, and computationally effective manner. After extensive analysis of a typical baseball game it was determined that a baseball game is usually captured by cameras positioned at fixed locations around the baseball field, with each camera typically capable of panning, tilting, and zooming. Each play in a baseball game normally starts with the pitcher releasing the ball, such as toward the catcher or toward one of the basemen. Further, a pitching scene, in which the pitcher is about to throw the ball, is usually captured from a camera location behind the pitcher. This camera angle is typically used because it is easier to observe the movements of all of the parties involved (the pitcher, the batter, the catcher, and the umpire) from this viewpoint.

It is based upon these observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that a region of generally green color and a region of a generally brown color should be used. Without any such realizations, there would be no motivation to modify Kawashima et al. to include color characteristics.

Claim 13 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, wherein the start of the plurality of



segments is identified based upon detecting at least one region of a generally green color and at least one region of a generally brown color.

The applicant would respectfully suggest that section 2.1.3 of Kawashima et al. uses count of pixels whose intensity change in successive frames are larger than a threshold. There is no use of color disclosed by Kawashima et al. to identify the play itself.

Claims 14-18 depend from claim 13 and are patentable for the same reasons asserted for claim 13.

As previously described, it is based upon those observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that at least three regions of characterized colors. Without any such realizations, there would be no motivation to modify Kawashima et al. to include color characteristics.

Claim 19 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, wherein the start of the plurality of segments is identified based upon detecting at least three regions of horizontally oriented regions of colors, where each of the segments includes a plurality of frames of the video.

Claims 20-23 depend from claim 19 and are patentable for the same reasons asserted for claim 19.

The applicant would respectfully suggest that section 2.1.3 of Kawashima et al. uses count of pixels whose intensity change in successive frames are larger than a threshold. There is no use of color disclosed by Kawashima et al. to identify the play itself.

As previously described, it is based upon those observations regarding the characteristics of a baseball broadcast and realization that part of the problem is the inability of complex sophisticated models to be sufficiently robust and computationally efficient that at least three regions of characterized image. Without any such realizations, there would be no motivation to modify Kawashima et al. to include color characteristics.

Claim 24 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, wherein the start of the plurality of segments is identified based upon identifying at least three regions of sufficient spatial coherence and sufficient horizontal extent, where each of the segments includes a plurality of frames of the video.

Claims 25-26 depend from claim 24 and are patentable for the same reasons asserted for claim 24.

The applicant would respectfully suggest that section 2.1.3 of Kawashima et al. uses count of pixels whose intensity change in successive frames are larger than a threshold. There is no use of color disclosed by Kawashima et al. to identify the play itself.

The Examiner notes that Kawashima does not explicitly disclose verifying that the first single frame is an appropriate start of the segment based upon processing of another single frame temporally relevant to the first single frame.

Claim 27 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, wherein the identifying for at least one of the segments includes detecting the start of the segment based upon processing of a first single frame of the video independently of other frames, where each of the segments includes a

plurality of frames of the video, and verifying that the first single frame is an appropriate start of the segment based upon processing of another single frame temporally relevant to the first single frame independently of other frames.

The applicant would note that Kawashima et al. disclose the determination of the first single frame and the verification of the first single frame using a temporal sequence of frames. There is no suggestion in Kawashima et al. to perform the analysis on a frame-by-frame manner.

Claim 28 depends from claim 27 and are patentable for the same reasons asserted for claim 27.

The Examiner notes that Kawashima et al. fail to disclose the identification of a segment that has a temporally sufficient short duration. Wang fail to suggest in the application of baseball summarization the desirability to remove segments of sufficiently short duration. The motivation to remove such segments is that it is likely to be a false positive.

Claim 39 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, where each of the segments includes a plurality of frames of the video and each of the segments is identification as likely including a play of the baseball, and identifying at least one of the segments that has a temporally sufficiently short duration.

Claims 40-41 depend from claim 39 and are patentable for the same reasons asserted for claim 39.

The Examiner notes that Kawashima et al. fail to disclose the identification of a commercials that is temporally close to a commercial. Wang fail to suggest in the application of baseball summarization the desirability to remove segments sufficiently close to a commercial.

Claim 44 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video wherein each of the segments includes a play of baseball, detecting a commercial within the video, and creating a summarization of the video by including the plurality of segments, where the summarization includes fewer frames than the video, where the plurality of segments of the video sufficiently temporally close to the commercial are free from being included within the summarization.

Claim 45 depends from claim 44 and are patentable for the same reasons asserted for claim 44.

The Examiner notes that Kawashima et al. fail to disclose the use of textures. The applicant would respectfully suggest that Kawashima et al. uses an intensity comparison technique in contrast to an analysis of the regions or motion aspects of the image, such as texture.

Claim 51 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying wherein the start of the plurality of segments is identified based upon detecting different textures.

Claim 52 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the video, wherein the start of the plurality of segments is identified based upon the video being free from substantial translational motion.

Claim 53 patentably distinguishes over Kawashima et al. in view of Wang et al. by claiming identifying a plurality of segments of the baseball video, wherein the identifying for the end of at least one of the segments is based upon detecting a scene change, where the scene change is not used to identify the end of the at least one of the segments if the temporally

relevant portion of the at least one the segment includes substantial horizontal motion, where each of the segments includes a plurality of frames of the baseball video.

The Examiner rejected claim 45 as being obvious over Kawashima et al. in view of Pan.

Claim 45 patentably distinguishes over the references of record by claiming identifying a plurality of segments of the video wherein each of the segments includes a play of baseball in a manner free from using a model sequence of frames to determine the approximate frame that the play starts, wherein the segments include full-speed plays and slow motion plays of the full-speed plays.

The Examiner rejected claims 36-38 and 40-41 as being unpatentable over Kawashima et al. in view of Wang et al. and further in view of Rui et al.

Claim 36 patentably distinguishes over Kawashima et al. in view of Wang et al. and further in view of Rui et al. by claiming identifying a plurality of segments of the video where each of the segments is identified as a play, where each of the segments includes a plurality of frames of the video, and identifying a plurality of the segments that are temporally separated by a sufficiently short duration.

There is no suggestion in these references to concatenate a plurality of segments each of which are identified as a play.

Claims 37-38 depend from claim 36 and are patentable for the same reasons asserted for claim 36.

The Examiner rejected claims 49 and 50 as being anticipated by Rui et al in view of Wang.

Appl. No. 09/934,004  
Amdt. dated December 20, 2004  
Reply to Office action of May 21, 2004

Rui et al. disclose the identification of highlights using audio-track features alone without relying on expensive-to-compute video track features. See, Rui, introduction, column 2.

Claim 49 patentably distinguishes over Rui et al. in view of Wnag by claiming identifying a plurality of segments of the video, where the identifying is based upon analysis of frames of the video, wherein each of the segments includes a play of baseball.

Claim 50 depends from claim 49 and are patentable for the same reasons asserted for claim 49.

All the pending claims should now be in a form suitable for allowance.

Respectfully submitted,

CHERNOFF, VILHAUER, MCCLUNG & STENZEL, LLP

A handwritten signature in black ink, appearing to read 'Kevin L. Russell', is written over a horizontal line.

Kevin L. Russell  
Reg. No. 38,292  
Tel No.: (503) 227-5631